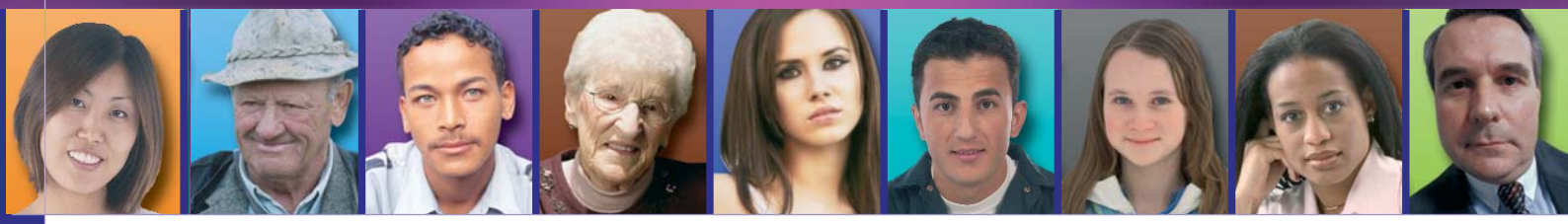
Making Chronic Conditions Count:

Hypertension
Stroke
Coronary Heart Disease
Diabetes



A systematic approach to estimating and forecasting population prevalence on the island of Ireland

Technical Supplement

Association of Public Health Observatories
The Centre of Excellence for Public Health Northern Ireland
HRB Health Research Board
Centre for Health and Diet Research

Making Chronic Conditions Count:

Hypertension
Stroke
Coronary Heart Disease
Diabetes

A systematic approach to estimating and forecasting population prevalence on the island of Ireland

Technical Supplement

Authors: Steve Barron, Kevin P Balanda, Lorraine Fahy

February 2010

A copy of the technical supplement, the full report and the executive summary can be downloaded from
http://www.inispho.org/publications/makingchronicconditionscounttechnicalsupplement
http://www.inispho.org/publications/makingchronicconditionscount
http://www.inispho.org/publications/makingchronicconditionscountexecutivesummary

Making Chronic Conditions Count: Hypertension, Stroke, Coronary Heart Disease, Diabetes. A systematic approach to estimating and forecasting population prevalence on the island of Ireland. Technical Supplement.

Published by the Institute of Public Health in Ireland.

Prepared by:
Steve Barron, Kevin P Balanda, Lorraine Fahy (Institute of Public Health in Ireland)

To be cited as:
Barron, S., Balanda, K.P., Fahy, L. *Making Chronic Conditions Count: Hypertension, Stroke, Coronary Heart Disease, Diabetes. A systematic approach to estimating and forecasting population prevalence on the island of Ireland. Technical Supplement.* Dublin: Institute of Public Health in Ireland, 2010.

The Institute would like to thank Mr John Hughes, Centre of Excellence for Public Health (Northern Ireland), Queen's University Belfast for his comments on earlier drafts of this document.

ISBN 978-0-9559598-6-8

Design by Slick Fish Design

For further copies of the technical supplement, the full report and the executive summary please contact:

Ireland and Northern Ireland's Population Health Observatory (INIsPHO),
The Institute of Public Health in Ireland (IPH).

5th Floor	Forestview
Bishop's Square	Purdy's Lane
Redmond's Hill	Belfast
Dublin 2	BT8 7ZX
Ireland	Northern Ireland
Tel: +353 1 478 6300	Tel: +44 2890 648494
Fax: +353 1 478 6319	Fax: +44 2890 646604

Email: info@inispho.org

A copy of the technical supplement, the full report and the executive summary can be downloaded from
http://www.inispho.org/publications/makingchronicconditionscounttechnicalsupplement
http://www.inispho.org/publications/makingchronicconditionscount
http://www.inispho.org/publications/makingchronicconditionscountexecutivesummary

Contents

Introduction

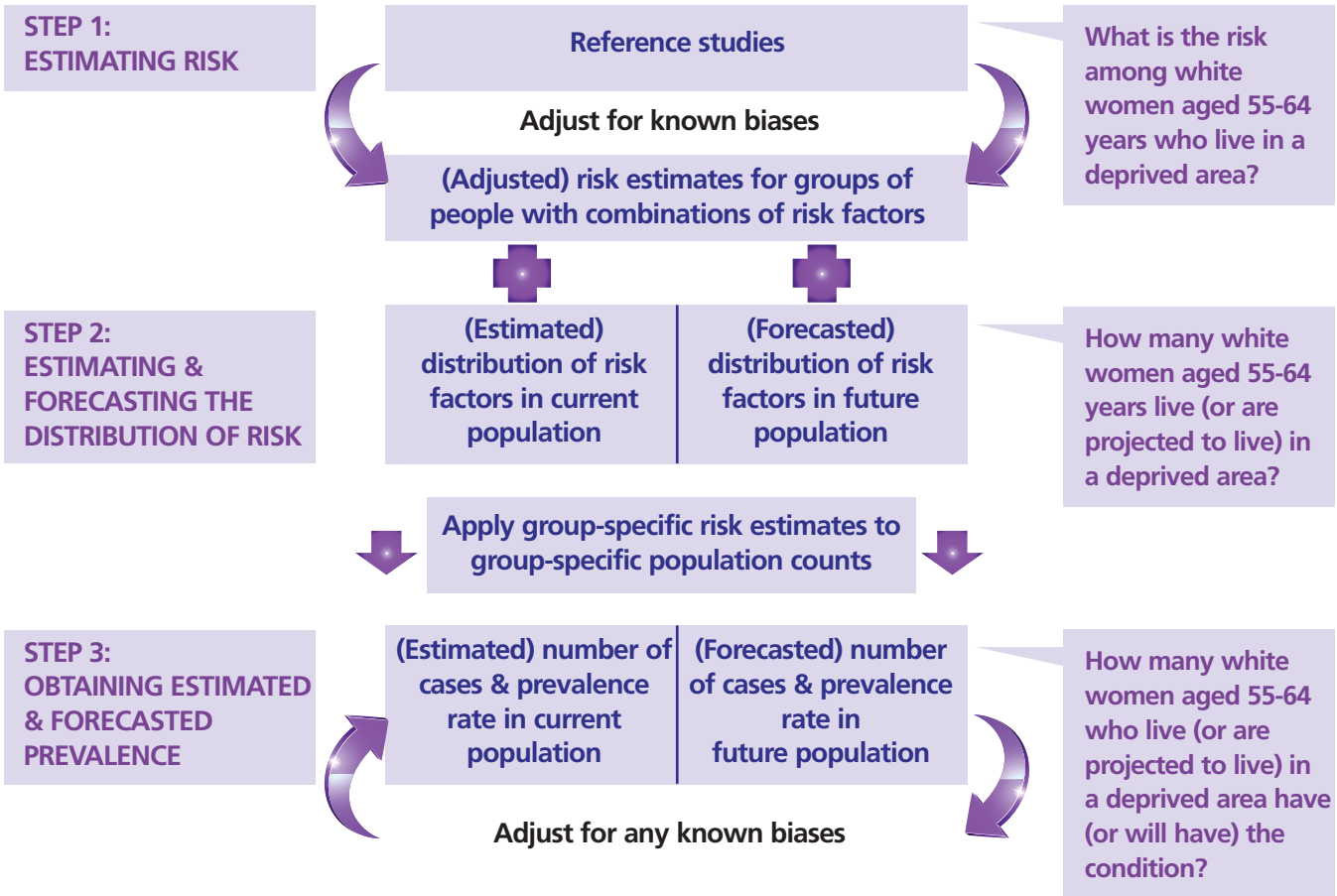
The Association of Public Health Observatories (APHO) (www.apho.org.uk) and its partners[1] were commissioned by the English Department of Health to develop models to estimate and forecast the population prevalence of a number of chronic conditions.

Ireland and Northern Ireland's Population Health Observatory (INIsPHO) in the Institute of Public Health in Ireland (IPH) has adapted the models for hypertension, angina and heart attack (coronary heart disease; CHD), stroke and diabetes to the island of Ireland. This technical supplement describes how the models were applied.

How the Models Work

Each model involves three simple steps that are described below:

Figure 1: How the models work



[1] The diabetes model was developed in collaboration with Brent PCT and the University of Sheffield, well before the models for the other conditions.

STEP 1: ESTIMATING RISK

Reference studies were used to reliably estimate the risk that an adult[2] with a particular combination of risk factors will have the condition. For example, what is the risk among white women aged 55-64 years living in a deprived area? In the case of diabetes, some additional adjustment (based on obesity rates) for known biases in the reference studies was also made.

STEP 2: ESTIMATING AND FORECASTING THE DISTRIBUTION OF RISK

The next step was to calculate the number of adults with these particular combinations of risk factors in the current and future population across the island. For example, how many white women aged 55-64 years live (or are projected to live) in a deprived area in 2007, 2015 and 2020?

STEP 3: OBTAINING ESTIMATED AND FORECASTED PREVALENCE

Group-specific risk estimates were then applied to corresponding group-specific population counts to estimate and forecast the number of adults with the condition and the population prevalence rate of the condition. For example, how many white women aged 55-64 years living in a deprived area have (or will have) the condition? In the case of diabetes, some additional adjustment for known biases was also made.

Step 1: Estimating Risk

Table 1 describes the definition, the reference studies and the risk factors used in the models for each condition.

Table 1: Definition, reference studies and risk factors for each condition

Condition	Definition	Reference studies	Risk factors
Hypertension	Measured systolic blood pressure (SBP) ≥ 140mmHg and/or measured diastolic blood pressure (DBP) ≥ 90mmHg or taking medicine prescribed for high blood pressure*.	Health Survey for England (2003 and 2004 combined).	Age, Sex, Ethnicity, Area deprivation.
Angina or heart attack (CHD)	Answered YES to the question 'Ever told by a doctor that you have angina or have had a heart attack?'	Health Survey for England (2003 and 2004 combined).	Age, Sex, Ethnicity, Area deprivation, Smoking.
Stroke	Answered YES to the question 'Ever told by a doctor that you have had a stroke?'	Health Survey for England (2003 and 2004 combined).	Age, Sex, Area deprivation, Smoking.
Diabetes (Type 1 and Type 2 combined)	WHO diagnostic criteria (1985) based on Glucose Tolerance Test.	Coventry Diabetes Study (Simmons, Williams and Powell, 1991). London-Brent Study (Chaturvedi, McKeigue and Marmot, 1993). Welsh Study (Harvey, Craney and Kelly, 2002).	Age, Sex, Ethnicity, Area deprivation, Obesity.

* Being hypertensive includes controlled, uncontrolled and untreated hypertension.

[2] Throughout most of this report, an 'adult' is aged 16 years or over. This applies to each condition except diabetes where an 'adult' is aged 20 years or over.

HYPERTENSION, CORONARY HEART DISEASE AND STROKE

Reference studies for hypertension, CHD and stroke consisted of analyses of the combined data from the Health Survey for England (HSfE) 2003 and 2004. APHO created a dataset for each condition that included all known risk factors that were available in HSfE. These datasets were used to create a logistic regression model for the conditions using stepwise selection of variables. However, not all the variables included in these 'complete' models were available at local geographical levels. 'Local' logistic regression models that included the risk factors from the 'complete' models that were available at local levels were then created. Appendix 1 shows the variables that were included in the 'complete' and the 'local' model for each condition.

Risk of disease was estimated as follows. The logistic regression models were used to derive the estimated odds of disease for all combinations of risk factors in the local model. Estimated odds were converted to estimated risks using the formula: risk = odds / (1+odds).

APHO validated the 'complete' and 'local' models in two ways. Firstly, a Receiver Operating Characteristics (ROC) curve was created for each model. Secondly, the disease status observed in the dataset was compared with the disease status predicted by the model. While each of the three 'complete' models performed slightly better than the corresponding 'local' model, all three 'local' models performed well in classifying disease status. Appendix 2 provides details of model validation.

APHO used the 'local' models to estimate and forecast population prevalence in Local Authorities and Primary Care Trusts in England. INIsPHO used the APHO 'local' models to estimate and forecast population prevalence in Local Health Offices (LHO) in the Republic of Ireland and Local Government Districts (LGD) in Northern Ireland.

DIABETES

The model for diabetes was based on three reference studies: the Coventry Diabetes Study (Simmons, Williams and Powell, 1991), the London-Brent Study (Chaturvedi, McKeigue and Marmot, 1993) and the Welsh Study (Harvey, Craney and Kelly, 2002).

Step 2: Estimating and Forecasting the Distribution of Risk

The distribution of risk factors in LHOs in the Republic of Ireland and LGDs in Northern Ireland were estimated for the years 2007, 2015 and 2020. If risk factor data were not available at LHO or LGD level the distribution of risk factors at a higher geographical level (where the risk factor data were available) was applied to the LHOs and/or LGDs within that area.

Republic of Ireland

SEX AND AGE

Population estimates for 2007 and population projections for 2015 and 2020 were based on the usually resident population. The Central Statistics Office did not identify a preferred population projection scenario among the scenarios they prepared. Only two scenarios were produced at sub-national level and the 'traditional' variant of the 'M2F1' scenario was used in this study. This scenario assumes fertility is fixed at 2006 level, moderately positive but declining international migration, and a traditional pattern of internal migration.

Population estimates and projections were not available for LHOs but were available for eight Regional Authorities. Age-sex specific changes in population from Census 2006 to 2007 (estimates), 2015 and 2020 (both projections) were calculated for each Regional Authority. These Regional Authority adjustment factors were applied to Census 2006 LHO data. For this we assumed that age-sex specific changes at Regional Authority level apply to each of the LHOs within that Regional Authority.

ETHNICITY

The ethnicity classifications used in the APHO models were based on the ethnic groups recorded in Census of England and Wales 2001. These ethnic groups are different from the ethnic groups recorded in the Republic of Ireland's Census 2006. The Republic of Ireland's ethnic groups were mapped to the English and Welsh ethnic groups as closely as possible. Tables 2 and 3 show the correspondences between ethnic groups in each country for the relevant conditions. Note that ethnicity was not included as a risk factor in the stroke model.

Table 2: Classifications of ethnic background used in the angina and heart attack (CHD) model and the hypertension model

Condition	Ethnic classification		
	Original APHO model England and Wales Census 2001	**Northern Ireland Census 2001**	**Republic of Ireland Census 2006**
Angina and heart attack (CHD); Hypertension	White • British • Irish • Other White	White • White • Irish Traveller	White • Irish • Irish Traveller • Other White • Not stated
	Mixed • White & Black Caribbean • White & Black African • White & Asian • Other Mixed	Mixed • Mixed	Mixed • Other (including Mixed)
	Asian • Indian • Pakistani • Bangladeshi • Other Asian	Asian • Indian • Pakistani • Bangladeshi	Asian • Asian (excluding Chinese)
	Black • Caribbean • African • Other Black	Black • Caribbean • African • Other Black	Black • African • Other Black
	Other • Chinese • Other	Other • Chinese • Other	Other • Chinese

Table 3: Classification of ethnic background used in the diabetes model

Condition	Ethnic classification		
	Original APHO model England and Wales Census 2001	**Northern Ireland Census 2001**	**Republic of Ireland Census 2006**
Diabetes	White and Mixed • British • Irish • Other White • White & Black Caribbean • White & Black African • White & Asian • Other Mixed	White and Mixed • White • Irish Traveller • Mixed	White and Mixed • Irish • Irish Traveller • Other White • Not stated • Other (including Mixed)
	Asian • Indian • Pakistani • Bangladeshi • Other Asian	Asian • Indian • Pakistani • Bangladeshi	Asian • Asian (excluding Chinese)
	Black • Caribbean • African • Other Black	Black • Caribbean • African • Other Black	Black • African • Other Black
	Other • Chinese • Other	Other • Chinese • Other	Other • Chinese

In the Republic of Ireland there were two misclassification issues:

- The original models for CHD, hypertension and diabetes provided separate estimates of the effect of a 'mixed' ethnic background and of an 'other' ethnic background. In the Republic of Ireland, people of a 'mixed' ethnic background are not recorded separately from people of an 'other' ethnic background. This meant that either people of a 'mixed' or people of an 'other' ethnic background had to be misclassified in the model. Northern Ireland's Census 2001 data showed that, amongst people aged 16 years or more, there were 1.5 times more people of a 'mixed' ethnic background than an 'other' ethnic background. It was assumed that the Republic of Ireland had a similar distribution and that classifying the 'Other (including Mixed)' group as 'mixed' rather than 'other' would result in less misclassification.

- It was assumed that people whose ethnicity was 'not stated' (1.4% of the reference population) were of a 'white' ethnic background. This misclassified people who did not state their ethnic background and were not of a 'white' ethnic background. With almost 94% of people in the Republic of Ireland aged 16 years or more being of a 'white' ethnic background in 2006, misclassification is likely to be low.

Population estimates and projections were not available by ethnicity. Age-sex breakdowns of LHO level ethnicity data were not available from Census 2006 in the Republic of Ireland due to disclosure concerns. Age-sex breakdowns of Regional Authority-level ethnicity data from Census 2006 were applied to the age-sex breakdowns of LHO-level population estimates and projections calculated in **SEX AND AGE** above. For this we assumed that the age-sex-ethnic distribution at Regional Authority level applies to each of the LHOs within that Regional Authority and that the ethnic distribution has not changed (and will not change) since Census 2006.

The age groups used in the diabetes model were different from those used in the models for the other conditions. The diabetes model's age groups led to disclosure concerns about age-sex breakdowns of Regional Authority-level ethnicity data from Census 2006. Therefore, the national age-sex-ethnic profile from Census 2006 was applied to the Republic of Ireland population estimates and projections. For this we assumed that the national age-sex-ethnic distribution applies to all LHOs and that the ethnic distribution has not changed (and will not change) since Census 2006.

DEPRIVATION

Local area deprivation scores for the Republic of Ireland were taken from New Measures of Deprivation for the Republic of Ireland (Haase and Pratschke, 2008; see Appendix 3 for details). It was assumed that an area's deprivation band would not change over time. The way in which local area deprivation scores were incorporated into the diabetes model is different than the way it was done in the models for the other conditions. See **STEP 3: OBTAINING ESTIMATED AND FORECASTED PREVALENCE** for details.

For the CHD, hypertension and stroke models in the Republic of Ireland, five deprivation bands were created by ordering the deprivation scores for all the Electoral Divisions (ED) and identifying cut-off scores that created five bands that included approximately equal numbers of EDs. An LHO's deprivation score was calculated as the population weighted average of the deprivation scores of the EDs within that LHO. These LHO deprivation scores were assigned to the deprivation bands that

had been based on ED deprivation scores. Therefore, there was not an equal number of LHOs within each band.

SMOKING

National age-sex specific proportions and Regional Authority-level overall proportions of cigarette smokers, ex-smokers and people who never smoked were available from SLÁN 2007. To estimate the smoking profile at sub-national level, the national age-sex specific estimates were adjusted by how much higher or lower the Regional Authority overall smoking prevalence rate was than the national overall smoking prevalence rate (see **METHOD TO ESTIMATE SMOKING PREVALENCE** below). It was assumed that the smoking profile at Regional Authority level applied to each of the LHOs within that Regional Authority. Smoking prevalence for people aged 16 and 17 years were not available and it was assumed that the smoking prevalence of people aged 18-24 years applied to people aged 16-24 years.

METHOD TO ESTIMATE SMOKING PREVALENCE

National smoking data were adjusted to estimate smoking prevalence in sub-national areas as follows:

- Sub-national proportion of smokers in age-sex group = national prevalence of smoking in age-sex group * sub-national overall smoking prevalence / national overall smoking prevalence
  S_as_sn = (S_as_n) * (S_sn) / (S_n)

- Sub-national proportion of ex-smokers in age-sex group was not adjusted
  E_as_sn = E_as_n

- Local proportion of never-smokers in age-sex group = 1 – (proportion of ex-smokers in age-sex group) – (local proportion of smokers in age-sex group)
  N_as_sn = 1 – (E_as_sn) – (S_as_sn)

Where:
S = proportion of population who were smokers
E = proportion of population who were ex-smokers
N = proportion of population who had never smoked
sn = sub-national (ie Regional Authorities in the Republic of Ireland and Health and Social Care Trusts in Northern Ireland)
n = national
as = age-sex specific

This approach assumed that:

- The proportions of smokers, ex-smokers and people who never smoked were the same across ethnic groups.
- The proportion of ex-smokers in each age-sex group was the same in all areas. This assumes that the number of people who never smoked increases as the number of smokers decreases. Regional analysis of the relationship between prevalence of smokers and ex-smokers in the HSfE showed no systematic relationship and it was decided that the ex-smoking rate should not be locally adjusted.

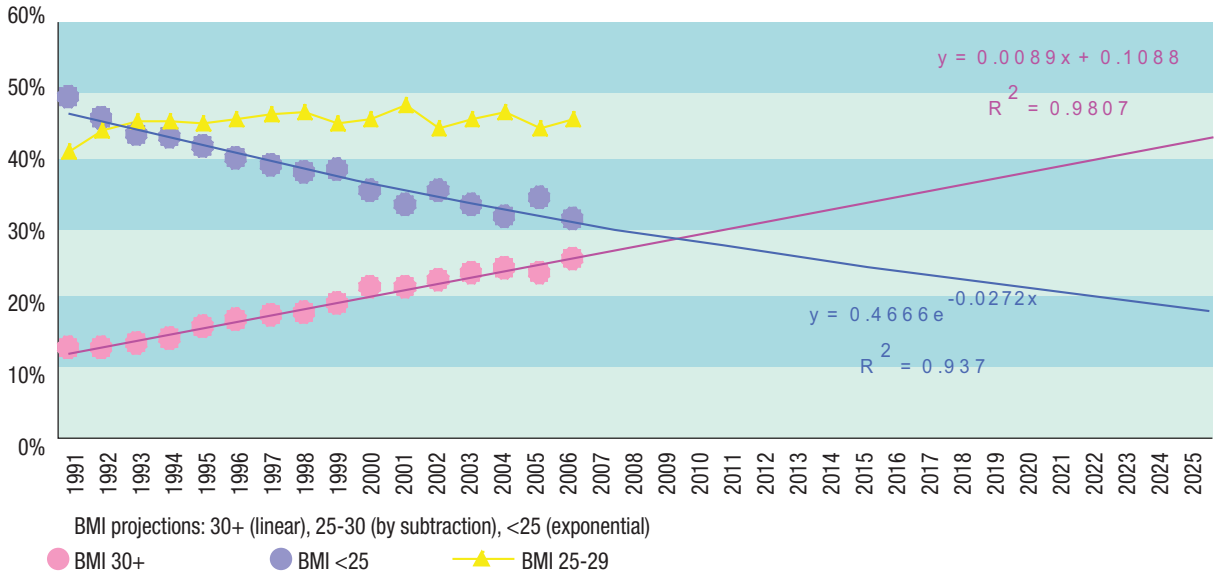
The APHO models used local synthetic estimates of smoking prevalence rather than direct estimates based on HSfE data. The synthetic estimates give the smoking prevalence that would be expected in an area given its population's characteristics as measured by census and administrative data. The smoking data used in Northern Ireland and the Republic of Ireland represent direct survey-based estimates.

Future changes in smoking prevalence were not taken into account in the prevalence forecasts. This was because of the uncertainty associated with predictions of smoking prevalence, and the lag time between smoking cessation and changes in health status.

OBESITY

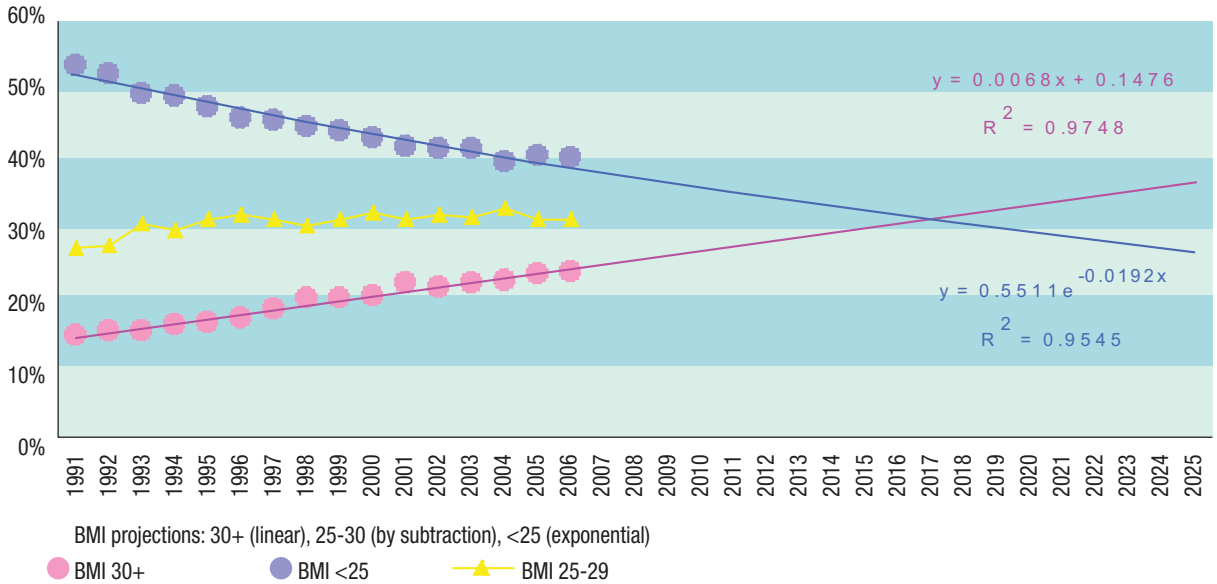
Obesity is included as a risk factor for Type 2 diabetes in the diabetes model. The population prevalence estimates from the original reference studies were adjusted to reflect the changes in obesity prevalence since the time of the reference studies to 2007, 2015 and 2020. The adjustment for 2007 was based on BMI data from HSfE 2006 as this was the most recently available HSfE data. The adjustments for 2015 and 2020 were based on regression analysis of HSfE data from 1991-2006 that forecast the future distribution of obesity in 2015 and 2020. The regression analysis used the 'most realistic' scenario from *Making Diabetes Count – What does the future hold?* (Jordan et al, 2007), which assumed that obesity rates will rise linearly and underweight/normal rates will slow exponentially between 2006 and 2020. The adjustment factors were calculated separately for males and females to reflect that obesity rates are rising more steeply in males than females. Figures 2 and 3 below show the regression models used to predict obesity levels up to 2025 in males and females respectively.

Figure 2: Projecting male BMI distribution to 2025



Source: Yorkshire and Humber Public Health Observatory (2008)

Figure 3: Projecting female BMI distribution to 2025



Source: Yorkshire and Humber Public Health Observatory (2008)

Northern Ireland

SEX AND AGE

Population estimates for 2007 and principal 2006-based population projections for 2015 and 2020 were based on the usually resident population. Population estimates and projections were available by age, sex and LGD.

ETHNICITY

The ethnicity classification used in the APHO models were based on the ethnic groups recorded in Census of England and Wales 2001. These ethnic groups are different to the ethnic groups recorded in Northern Ireland's Census 2001. Northern Ireland's ethnic groups were mapped to provide the best match for England and Wales' classification. Generally there was a good match between the classification of ethnic groups in England and Wales and the remapped ethnic groups in Northern Ireland. Tables 2 and 3 show the classification of ethnic groups for each country for the relevant conditions. Note that ethnicity was not included as a risk factor in the stroke model.

Population estimates and projections were not available by ethnicity. Age-sex breakdowns of LGD-level ethnic proportions from Census 2001 were applied to corresponding age-sex LGD-level population estimates and projections. For this we assumed that the ethnic distribution has not changed (and will not change) since Census 2001.

DEPRIVATION

Local area deprivation scores for Northern Ireland were taken from the Northern Ireland Multiple Deprivation Measure 2005 (see Appendix 3 for details). It was assumed that an area's deprivation band would not change over time. The way in which local area deprivation scores were incorporated into the diabetes model is different than the way it was done in the models for the other conditions. See **STEP 3: OBTAINING ESTIMATED AND FORECASTED PREVALENCE** for details.

Similar to the models for CHD, hypertension and stroke in the Republic of Ireland, five deprivation bands were created by ordering the deprivation scores for all the Super Output Areas (SOA) and identifying cut-off scores that created five bands that included approximately equal numbers of SOAs. An LGD's deprivation score was calculated as the population weighted average of the deprivation scores of the SOAs within that LGD. These LGD deprivation scores were assigned to the deprivation bands that had been based on SOA deprivation scores. There was not an equal number of LGDs within each band and no LGD score was within the least deprived national quintile of SOA scores. Therefore, there were only four deprivation bands in Northern Ireland.

SMOKING

National age-sex specific proportions and Health and Social Care (HSC) Trust-level overall proportions of cigarette smokers, ex-smokers and people who never smoked were available from Continuous Household Survey 2007/2008. To estimate the smoking profile at sub-national

level, the national age-sex specific estimates were adjusted by how much higher or lower the HSC Trust overall smoking prevalence rate was than the national overall smoking prevalence rate (see **METHOD TO ESTIMATE SMOKING PREVALENCE** above). It was assumed that the smoking profile at HSC Trust level applied to each of the LGDs within that HSC Trust.

OBESITY

The obesity adjustments for the diabetes model were calculated in the same way as the Republic of Ireland. See **REPUBLIC OF IRELAND: OBESITY** for details.

Step 3: Obtaining Estimated and Forecasted Prevalence

HYPERTENSION, CORONARY HEART DISEASE AND STROKE

The group-specific risk estimates calculated in Step 1 were applied to group-specific population counts/projections that were estimated/forecasted in Step 2. Prevalence estimates for 2007 and prevalence forecasts for 2015 and 2020 were produced for each LHO in the Republic of Ireland and each LGD in Northern Ireland. Figures were broken down by age, sex, ethnicity (where applicable) and local area deprivation.

DIABETES (TYPE 1 AND TYPE 2)

Initial diabetes prevalence estimates for 2007 and prevalence forecasts for 2015 and 2020 were produced for each LHO in the Republic of Ireland and each LGD in Northern Ireland. Figures were broken down by age, sex and ethnicity.

Only Type 2 diabetes estimates and forecasts were then adjusted to account for the effect of local area deprivation. The cases of Type 2 diabetes were redistributed so that sub-national variation in prevalence reflected sub-national variation in deprivation but the overall number of diabetes cases in the Republic of Ireland and Northern Ireland was not adjusted. The redistribution was achieved by applying an adjustment factor to prevalence estimates for each LHO and LGD. LHO adjustment factors were calculated using Electoral Divisions (ED) and LGD adjustment factors were calculated using Super Output Areas (SOA).

Calculation of the adjustment factors involved the following steps:

- EDs/SOAs were ordered from most deprived to least deprived. The ordered data for each country were split into five deprivation bands of approximately equal population size.
- The percentage of LHO/LGD population living in each national deprivation band was calculated.
- The risks of Type 2 diabetes associated with each national deprivation band were taken from England's National Diabetes Audit (NDA) (see Table 4). (Health and Social Care Information Centre, 2007).
- These deprivation-specific risks were applied to the deprivation profile of each LHO/LGD to obtain the number of diabetes cases that would be expected.
- The number of cases expected in an area divided by the number expected nationally in the Republic of Ireland or Northern Ireland was taken as the adjustment factor for that area.

Table 4: Registered Type 2 diabetes prevalence from National Diabetes Audit

	Most deprived				Least deprived
Deprivation bands	**1**	**2**	**3**	**4**	**5**
Registered diabetes prevalence	3.78%	3.50%	3.25%	3.01%	2.58%

The adjustment for deprivation may be conservative due to the following reasons (YHPHO, 2008):

- The adjustment assumed that the gradient of registered Type 2 diabetes across deprivation bands was the same as the gradient of the total (diagnosed plus undiagnosed) Type 2 diabetes across deprivation bands. The rate of diagnosis may be lower among persons living in more deprived areas, hence the gradient of registered Type 2 diabetes across deprivation bands may be an under-estimate of the gradient of total Type 2 diabetes across deprivation bands.

- The adjustment assumed that each area had the same ethnic profile. However, people from minority ethnic groups may be more likely to live in deprived areas.

- The diabetes prevalence figures from the NDA were not age-standardised. Since less deprived areas tend to have older populations, the age-standardised deprivation gradient of diabetes prevalence is likely to be steeper than shown by the crude rates in the model.

References

Balanda, K.P., Fahy, L., Jordan, A., McArdle, E. (2006). *Making Diabetes Count: A systematic approach to estimating and forecasting population prevalence on the island of Ireland in 2005.* Dublin: Institute of Public Health in Ireland. Available at: http://www.inispho.org/publications/makingdiabetescountasystematicapproachtoestimatingpopulationprevalenceontheislandofirel

Balanda, K.P., Barron, S., Fahy, L., McLaughlin, A. (2010). *Making Chronic Condition Count: Hypertension, Coronary Heart Disease, Stroke, Diabetes. A systematic approach to estimating and forecasting population prevalence on the island of Ireland.* Dublin: Institute of Public Health in Ireland, 2010. Available at: http://www.inispho.org/publications/makingchronicconditionscount

Balanda, K.P., Barron, S. and Fahy, L. (2010). *Making Chronic Conditions Count: Hypertension, Coronary Heart Disease, Stroke, Diabetes.* A systematic approach to estimating and forecasting population prevalence on the island of Ireland. Executive Summary. Dublin: Institute of Public Health in Ireland. Available at: http://www.inispho.org/publications/makingchronicconditionscountexecutivesummary

Chaturvedi, N., McKeigue, P.M., Marmot, M.G. (1993). *Resting and ambulatory blood pressure differences in Afro-Caribbeans and Europeans.* Hypertension, Vol 22, pp 90-6. Available at: http://hyper.ahajournals.org/cgi/reprint/22/1/90

Haase, T. and Pratschke, J. (2008). *New Measures of Deprivation for the Republic of Ireland.* Dublin: Pobal. Available at: http://www.pobal.ie/WhatWeDo/Deprivation/Pages/InformationforBeneficiaries.aspx

Harvey, J.N., Craney, L., Kelly, D. (2002). *Estimation of the prevalence of diagnosed diabetes from primary and secondary care source data: comparison of record linkage with capture-recapture analysis.* Journal of Epidemiology & Community Health, Vol 56, pp 18-23. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1731996/pdf/v056p00018.pdf

Health and Social Care Information Centre (2007). *National Diabetes Audit: Key findings about the quality of care for people with diabetes in England and Wales. Report for the audit period 2005-2006.* Available at: http://www.ic.nhs.uk/services/national-clinical-audit-support-programme-ncasp/audit-reports/diabetes

Jordan, A., Graham, A., Balanda, K.P., Fahy, L. (2007). *Making Diabetes Count: What does the future hold? A systematic approach to forecasting population prevalence on the island of Ireland in 2010 and 2015.* Dublin: Institute of Public Health in Ireland. Available at: http://www.inispho.org/publications/makingdiabetescountwhatdoesthefuturehold

Northern Ireland Statistics & Research Agency (2005). *Northern Ireland Multiple Deprivation Measure 2005.* Available at: http://www.nisra.gov.uk/deprivation/nimdm_2005.htm

Simmons, D., Williams, D.R., Powell, M.J. (1991). *The Coventry Diabetes Study: Prevalence of Diabetes and Impaired Glucose Tolerance in Europids and Asians.* Quarterly Journal of Medicine, Vol 81, pp 1021-30. Available at: http://qjmed.oxfordjournals.org/cgi/content/abstract/81/3/1021

Yorkshire and Humber Public Health Observatory (2008). *PBS Diabetes Model Phase 3 Technical Briefing Document.* Available at: http://www.yhpho.org.uk/resource/view.aspx?RID=9906

APPENDIX 1:
Risk factors included in hypertension, coronary heart disease and stroke models

Two logistic regression models were created for each of hypertension, CHD and stroke. A 'complete' model was created using stepwise selection of known risk factors from HSfE data. However, not all the variables included in the 'complete' model were available at local geographical levels. APHO also created a 'local' logistic regression model for each condition that included the risk factors from the 'complete' model for which data was available at local levels.

Full details of APHO's prevalence modelling work can be accessed at www.apho.org.uk/resource/item.aspx?RID=48308.

Table 5: Variables in the 'complete' and 'local' hypertension models

'Complete' hypertension model's variables and levels	Included in the 'local' hypertension model?
Age 16-24 years 25-34 years 35-44 years 45-54 years 55-64 years 65-74 years 75+ years	Y
Sex Male Female	Y
Ethnic group White Mixed Black Asian Other	Y
Area deprivation National quintile 1 National quintile 2 National quintile 3 National quintile 4 National quintile 5 National quintiles of deprivation were based on Index of Multiple Deprivation 2004 scores at Lower Super Output Area level in England	Y
Body Mass Index Under 20 20 to 25 25 to 30 Over 30	N
Days in the previous 4 weeks with physical activity None Less than one One or two a week Three or four a week Five or more a week	N
Salt added at the table Generally add salt to food without tasting it first Taste the food, but then generally add salt Taste the food, but only occasionally add salt Rarely, or never, add salt at the table	N
Limiting long term illness Limiting long term illness Non limiting long term illness No long term illness	N
Highest educational qualification NVQ4/NVQ5/degree or equivalent Higher education below degree NVQ3/GCSE A level equivalent NVQ2/GCSE O level equivalent NVQ1/GCSE other grade equivalent Foreign educated/other No qualification Fulltime student	N

Table 6: Variables in the 'complete' and 'local' coronary heart disease (CHD) models

'Complete' CHD model's variables and levels	Included in the 'local' CHD model?
Age 16-24 years 25-34 years 35-44 years 45-54 years 55-64 years 65-74 years 75+ years	Y
Sex Male Female	Y
Ethnic group White Mixed Black Asian Other	Y
Area deprivation National quintile 1 National quintile 2 National quintile 3 National quintile 4 National quintile 5 National quintiles of deprivation were based on Index of Multiple Deprivation 2004 scores at Lower Super Output Area level in England	Y
Smoking Never smoked Used to smoke occasionally Used to smoke regularly Current smoker	Y
Body Mass Index Under 18.5 18.5 to 25 25 to 30 30 to 40 Over 40	N
Total cholesterol to HDL cholesterol ratio	N
Diabetes Yes No	N
Family history of cardiovascular disease Yes No	N
Limiting long term illness Limiting long term illness Non limiting long term illness No long term illness	N
WHO Rose questionnaire Angina and myocardial infarction Neither angina nor myocardial infarction Angina but not myocardial infarction Myocardial infarction but not angina	N

Table 6: Variables in the 'complete' and 'local' coronary heart disease (CHD) models (cont.)

'Complete' CHD model's variables and levels	Included in the 'local' CHD model?
Taking ACE inhibitor Yes No	N
Taking beta blocker Yes No	N
Taking calcium blocker Yes No	N
Taking diuretic Yes No	N
Taking other hypertension medication Yes No	N
Highest educational qualification NVQ4/NVQ5/degree or equivalent Higher education below degree NVQ3/GCSE A level equivalent NVQ2/GCSE O level equivalent NVQ1/GCSE other grade equivalent Foreign educated/other No qualification Fulltime student	N

APHO created two 'complete' CHD models. One 'complete' model included hypertension medication as a risk factor while the other 'complete' model did not. The 'complete' model including hypertension medication is documented here and in Appendix 2 as it has a higher validity as measured by the area under the Receiver Operating Characteristics Curve.

Table 7: Variables in the 'complete' and 'local' stroke models

'Complete' stroke model's variables and levels	Included in the 'local' stroke model?
Age 16-24 years 25-34 years 35-44 years 45-54 years 55-64 years 65-74 years 75+ years	Y
Sex Male Female	Y
Area deprivation National quintile 1 National quintile 2 National quintile 3 National quintile 4 National quintile 5 National quintiles of deprivation were based on Index of Multiple Deprivation 2004 scores at Lower Super Output Area level in England	Y
Smoking Never smoked Used to smoke occasionally Used to smoke regularly Current smoker	Y
Body Mass Index Under 20 20 to 25 25 to 30 Over 30	N
Ever had hypertension Yes No	N
Limiting long term illness Limiting long term illness Non limiting long term illness No long term illness	N
GHQ12 score 1 2 3 4 5 6 7 8 9 10 11 12	N

APHO created three 'complete' stroke models each with a different definition of hypertension as a risk factor: i) history of hypertension; ii) currently on hypertension medication; iii) clinical measurement at time of survey. The 'complete' model with history of hypertension is documented here and in Appendix 2 as it has the highest validity as measured by the area under the Receiver Operating Characteristics Curve.

APPENDIX 2:
Validation of hypertension, coronary heart disease and stroke models

Two logistic regression models were created for each of hypertension, CHD and stroke. A 'complete' model was created using stepwise selection of known risk factors from HSfE data. However, not all the variables included in the 'complete' model were available at local geographical levels. APHO also created a 'local' logistic regression model for each condition that included the risk factors from the 'complete' model for which data was available at local levels. Appendix 1 shows the variables that were included in the 'complete' and the 'local' model for each condition.

APHO validated the 'complete' and 'local' models in two ways. Firstly, a Receiver Operating Characteristics (ROC) curve was created for each model. Secondly, the disease status observed in the dataset was compared with the disease status predicted by the model.

RECEIVER OPERATING CHARACTERISTICS (ROC) CURVE

ROC analysis is a useful tool for evaluating the performance of diagnostic tests and for evaluating the accuracy of a statistical model that classifies people into one of two categories, eg diseased or non-diseased. A ROC curve is a plot of sensitivity on the y axis against (1-specificity) on the x axis for varying values of the threshold t. The area under the curve (AUC) is an overall summary of diagnostic accuracy. AUC equals 0.5 when the ROC curve corresponds to random chance and 1.0 for perfect accuracy. If both sensitivity and specificity are important in a diagnostic model, the optimal threshold of t would be 0.75, where sensitivity and specificity equal 0.77.

Table 8: Area under Receiver Operating Characteristics curve (AUC) for the 'complete' and 'local' hypertension, coronary heart disease and stroke models.

	Area under ROC curve (AUC)	
	'Complete' model	'Local' model
Hypertension	0.83	0.81
Coronary heart disease	0.92	0.84
Stroke	0.87	0.83

While each of the three 'complete' models performed slightly better than the corresponding 'local' model, all three 'local' models' AUC values are above the optimal value of 0.75.

PREDICTION OF DISEASE STATUS

Statistical models summarise the relationship in a dataset between risk factors and an outcome such as disease status. A model can predict the outcome for each observation in the dataset but, depending on how well the model fits the data, the outcome predicted by the model may not be the same as the outcome observed in the data. Valid statistical models have a good match between predicted and observed outcomes.

Table 9: Disease status as predicted by each model and disease status as observed in the data.

Condition	Model	Observed	Predicted: No	Predicted: Yes	Total
Hypertension	Complete	No	6,446	938	7,384
		Yes	1,424	1,878	3,302
		Total	7,870	2,816	10,686
	Local	No	6,886	1,129	8,015
		Yes	1,717	2,012	3,729
		Total	8,603	3,141	11,744
Coronary heart disease	Complete	No	3,883	60	3,943
		Yes	203	246	449
		Total	4,086	306	4,392
	Local	No	17,614	3	17,617
		Yes	1,098	1	1,099
		Total	18,712	4	18,716
Stroke	Complete	No	16,366	1	16,367
		Yes	302	0	302
		Total	16,668	1	16,669
	Local	No	20,740	0	20,740
		Yes	493	0	493
		Total	21,233	0	21,233

Shaded cells show the number of observations where predicted disease status is different to observed disease status.

The shaded cells show the number of observations where predicted disease status is different to observed disease status. While each of the three 'complete' models result in less misclassification than the corresponding 'local' model, all three 'local' models performed satisfactorily in predicting disease status.

APPENDIX 3:
Deprivation measures used from Northern Ireland and the Republic of Ireland

NORTHERN IRELAND MULTIPLE DEPRIVATION MEASURE 2005

This deprivation index consists of seven domains. Each domain consists of a number of indicators. The domains and indicators are shown in Table 10. Indicators were combined to give a score for each domain. Domain scores are weighted and combined to give an overall deprivation score for each area (see Northern Ireland Statistics and Research Agency (NISRA), 2005 for full details).

Table 10: Deprivation domains, domain weights, and indicators, Northern Ireland Multiple Deprivation Measure 2005

Domain	Domain weight	Indicators
Income	25%	• Adults and children in Income Support households (includes lone parents and Minimum Income Guarantee recipients) (2003, Source: DSD) • Adults and children in income based Job Seeker's Allowance households (2003, Source: DSD) • Adults and children in Working Families' Tax Credit households whose equivalised income (excluding housing benefits) is below 60% of median before housing costs (2003, Source: Inland Revenue and DSD) • Adults and children in Disabled Person's Tax Credit households whose equivalised income (excluding housing benefits) is below 60% of median before housing costs (2003, Source: Inland Revenue and DSD)
Employment	25%	• Unemployment claimant count (JUVOS) of women aged 18- 59 and men aged 18-64 averaged over 4 quarters (2003, Source: DETI) • Incapacity Benefit claimants women aged 18-59 and men aged 18-64 (2003, Source: DSD) • Severe Disablement Allowance claimants women aged 18-59 and men aged 18-64 (2003, Source: DSD) • Participants in New Deal for Young People (18-24 years) who are not included in the claimant count (2003, Source: DEL) • Participants in New Deal for 25+ who are not included in the claimant count (2003, Source: DEL) • Invalid Care Allowance claimants women aged 18-59 and men aged 18-64 (2003, Source: DSD)
Health Deprivation and Disability	15%	• Years of Potential Life Lost (1999 to 2003, Source: Mortality data, NISRA) • Comparative Illness and Disability Ratio (2003, Source: IS, AA, DLA, SDA, IB from DSD) • A combined measure of two indicators (i) individuals suffering from mood or anxiety disorders, based on prescribing (2003, Source: CSA) and (ii) suicides (1999 to 2003, Source: NISRA) • People registered as having cancer (excluding non-melanoma skin cancers) (1999 to 2002, Source: Northern Ireland Cancer Registry)

Table 10: Deprivation domains, domain weights, and indicators, Northern Ireland Multiple Deprivation Measure 2005 (cont.)

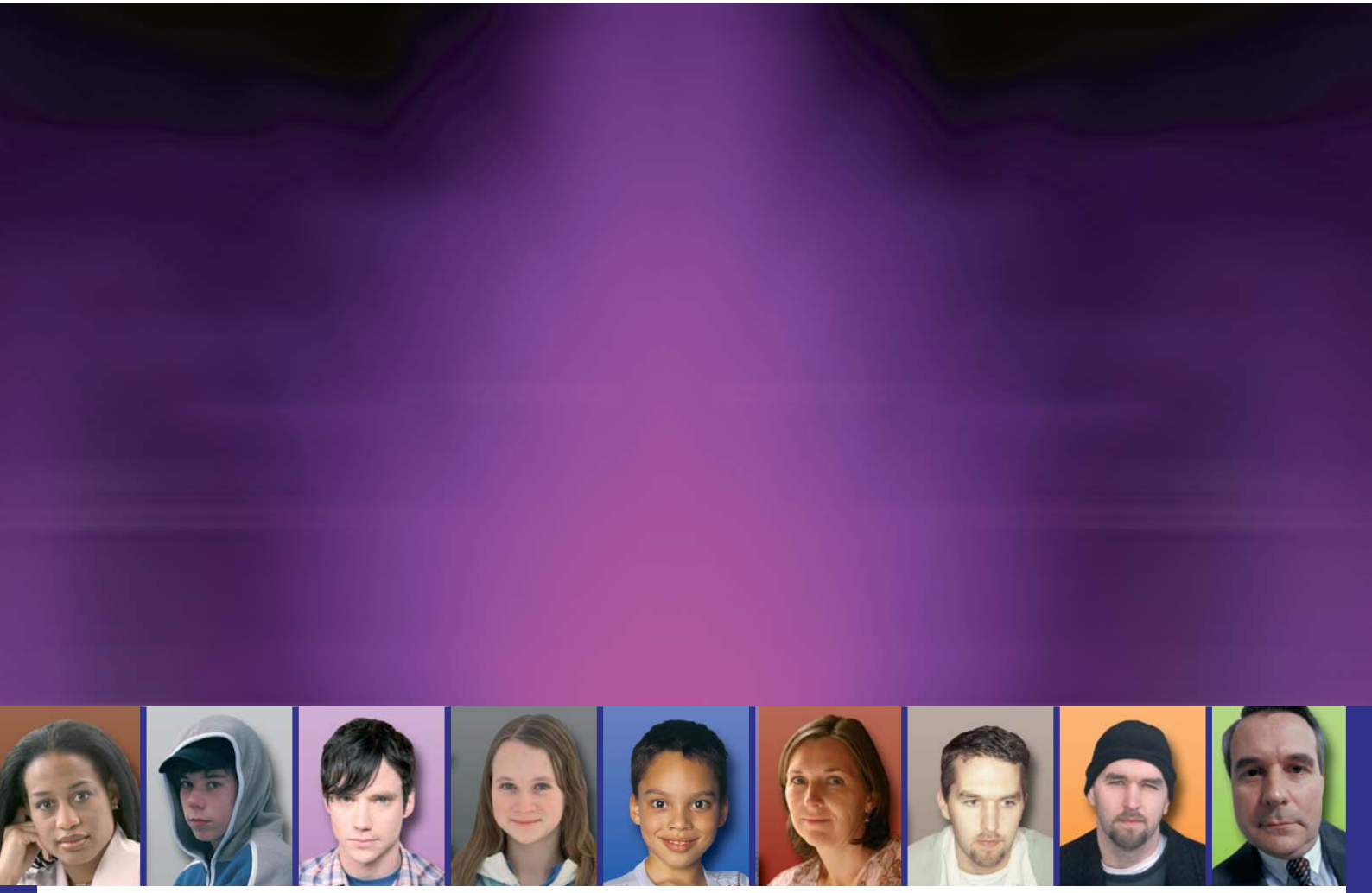
Domain	Domain weight	Indicators
Education, Skills and Training	15%	**Children/Young people** • GCSE/GNVQ points score (1999/2000 to 2001/2002, Source: School Leavers Survey, DE) • Key Stage 3 data (2002/2003, Source: DE) Note: Key Stage 3 assessment is based on formal tests taken by pupils at the end of KS3 (approximately age 14) in English (and Irish - in Irish medium schools/units), Mathematics and Science • Proportions of those leaving school aged 16 and not entering Further Education (1999/2000 to 2001/2002, Source: School Leavers Survey, DE) • Absenteeism at secondary level (all absences) (2001/2002 and 2002/2003 Source: SAER, DE) • Proportions of 17-20 year olds who have not successfully applied for Higher Education (1999/2000 to 2002/2003, Source: UCAS and DEL) • Proportions of Years 11 and 12 pupils not in a grammar school (2003, Source: School Census, DE) • Proportions of post primary pupils with Special Educational Needs in mainstream schools (2002/2003 School Census, Source DE) **Adults** • Proportions of working age adults (aged 25-59) in the area with no or low levels of qualification (2001, Source: Census, NISRA)
Proximity to Services	10%	• Road distance to a GP premises (2004, Source: CSA) • Road distance to an Accident and Emergency hospital (2004, Source: DHSSPS) • Road distance to a dentist (2004, Source: CSA) • Road distance to an optician (2004, Source: CSA) • Road distance to a pharmacist (2004, Source: CSA) • Road distance to a Job Centre or Jobs and Benefit office (2004, Source: DEL) • Road distance to a Post Office (2004, Source: Post Office Ltd) • Road distance to a food shop (2003, Source: Census of Employment) • Road distance to the centre of a settlement of 10,000 or more people (2004, Source: NISRA)
Living Environment	5%	**Housing quality** • SOA level housing stress (2001, Source: SDRC and NIHE, modelled NIHCS) • Houses without central heating (2001, Source: Census, NISRA) **Housing access** • Household overcrowding (2001, Source: Census, NISRA) • LGD level rate of acceptances under the homelessness provisions of the Housing (Northern Ireland) Order 1988 and the Housing (Northern Ireland) Order 2003, assigned to the constituent SOAs (2003, Source: NIHE) **Outdoor physical environment** • SOA level local area problem score (2001, Source: SDRC and NIHE, modelled NIHCS)
Crime and Disorder	5%	**Crime** • Violence, robbery and public order (April 2002 to March 2004, Source: PSNI) • Burglary (April 2002 to March 2004, Source: PSNI) • Vehicle theft (April 2002 to March 2004, Source: PSNI) • Criminal damage (April 2002 to March 2004, Source: PSNI) **Disorder** • Malicious and deliberate primary fires (April 2002 to March 2004, Source: NIFB) • Disturbances (April 2002 to March 2004, Source: PSNI)

New Measures of Deprivation in the Republic of Ireland

This deprivation index is based on Census 2006 and consists of three dimensions. These dimensions are linked to indicators from Census 2006 using Confirmatory Factor Analysis. The dimensions and indicators are shown in Table 11. Indicators were combined to give a score for each dimension. Indicator and dimension scores are combined to give and overall deprivation score for each area (see Haase and Pratschke (2008) for full details).

Table 11: Deprivation dimensions and indicators, New Measures of Deprivation in the Republic of Ireland

Dimension	Indicator
Demographic Profile	• The percentage increase in population over the previous five years • The percentage of population aged under 15 or over 64 years of age • The percentage of population with a primary school education only • The percentage of population with a third level education • The percentage of households with children aged under 15 years and headed by a single parent
Social Class Composition	• The percentage of population with a primary school education only • The percentage of population with a third level education • The percentage of households headed by professionals or managerial and technical employees, including farmers with 100 acres or more • The percentage of households headed by semi-skilled or unskilled manual workers, including farmers with less than 30 acres • The mean number of persons per room
Labour Market Situation	• The percentage of households headed by semi-skilled or unskilled manual workers, including farmers with less than 30 acres • The percentage of households with children aged under 15 years and headed by a single parent • The male unemployment rate • The female unemployment rate



The Institute of Public Health in Ireland

5th Floor
Bishop's Square
Redmond's Hill
Dublin 2
Tel: +353 1 478 6300
Fax: +353 1 478 6319

Forestview
Purdy's Lane
Belfast
BT8 7ZX
Tel: +44 2890 648494
Fax: +44 2890 646604

ISBN: 978-0-9559598-6-8

IPH

Design: www.slickfish.ie